

Strategic Plan for Research Data Services at the University of Minnesota Libraries

Final Report of the Research Data Services Strategic Planning Task Force Delivered March 13, 2018

Members of the Strategic Planning Task Force:

Lisa Johnston (Chair), Research Data Management &
Curation Lead, University Libraries

Shannon Farrell, Natural Resources Librarian

Philip Herold, Director for Agricultural, Biological,
and Environmental Sciences and Interim
Director for Physical Sciences and Engineering

Alicia Hofelich Mohr, Research Data Manager
(College of Liberal Arts, LATIS)

Shanda Hunt, Public Health Library Liaison & Data
Curation Specialist

Kelly Thompson, Metadata Analyst

Strategic Planning Task Force Sponsors:

Claire Stewart, Associate University Librarian for
Research & Learning

John Butler, Associate University Librarian for Data
& Technology

Janet Bishop, Associate University Librarian for
Content & Collections

Janice M. Jaguszewski, Associate University
Librarian and Director, Health Sciences
Libraries

Cite this report as:

University of Minnesota Libraries. (2019). *Strategic Plan for Research Data Services at the University of Minnesota Libraries*. Retrieved from the University of Minnesota Digital Conservancy, <http://hdl.handle.net/11299/203926>.

Table of Contents

Strategic Plan Summary	3
#1 Enhance Support Across the Research Workflow	3
#2 Provide Support for New/Emerging Needs	5
#3 Build on Successful Education/Training	6
#4 Facilitate Data Access	8
#5 Provide Services for Long-term Data Stewardship	9
#6 Align Research Data Services	11
Background	14
Methodology	15
Stage 1: Information Collection and Analysis	15
Stage 2: Stakeholder engagement around preliminary analysis	18
Stage 3: Finalize data services strategic plan	21
Results	22
Bibliography	22
Appendix A: Research Data Services 2017 “Themes”	25
Appendix B: Charge for University Libraries Research Data Services Strategic Planning Task Force	43

Executive Summary

Based on evidence of need from in-person campus stakeholder discussions in 2017 and an environmental scan of 24 reports (2004-2018) we identified six broad categories and 18 themes that are near-term priorities for Research Data Service (RDS) at the University of Minnesota Libraries.

#1 Enhance Support Across the Research Workflow

The University Libraries will enhance existing research data services offered across the entire research workflow.

Data Management

- Offer timely “rent a data manager” services
- Incorporate data management skills into the curriculum
- Support for archival and analog data management (beyond digital)

Reproducible Research

- Market tools for greater research transparency (e.g., OSF, DRUM)
- Pilot code run/review service
- Provide example workflows to ease adoption

Study Design / Research Planning

- Enhance IRB/participant agreement guidance for data sharing
- Align our DMP training with outreach by grant coordinators

Data Management Plans (DMPs)

- Develop active or machine- actionable DMP templates
- Systematically follow-up with U of M researchers to help them implement their DMP
- Outreach to masters and PhD students

#2 Provide Support for New & Emerging Needs

The University Libraries will recognize and respond to new and emerging data-related needs that our campus researchers face.

Data Visualization

- Provide access to a range of licensed data visualization tools
- Develop expertise in how to select and use data visualization tools; partner with others on campus that host visualization tools as appropriate

Data Use Agreements (DUA's)

- Consultation for DUAs for requesting existing data from an external source
- Writing DUAs for data the researcher/University produces to be shared to others

#3 Build on Successful Education/Training

The University Libraries will scale successful training programs in research data management to reach a broader audience.

Workshops

- Refine the successful “Data Management Bootcamp” to more targeted audiences

Course Integration

- Expand successful data management education in methods courses

#4 Facilitate Data Access

The University Libraries are exceptional providers of content in all formats, including data, and will enhance users’ access to our growing collection of research data.

Collection Development

- Establish an internal data purchasing/hosting procedures
- Communicate how purchasing decisions are made

Discovery and Access

- Create a “Data Registry” discovery layer showcase datasets that are of strategic importance to U of Minnesota and the state.

#5 Provide Services for Long-term Data Stewardship

The University Libraries will ensure accessibility and preservation of research data through curation, metadata, repositories, and other access and retrieval mechanisms to meet federal, state, sponsor, and University requirements.

Project Close Out

- Guidance and consultation around preparing data for sharing and archiving
- Coordinate with storage units who archive data around retention best practices
- Outreach to graduating students about managing data before they leave

Data Preservation

- Raise awareness for preservation of valuable data (i.e., unique, institutional assets)
- Work with campus stakeholders on data preservation to better support the research process

Data Repository and Curation

- Support “data curation in place” for data archived in non-library repositories
- Build workflows to better support data deposited to DRUM for peer review alongside journal article manuscript
- Create a staging platform for manipulating DRUM files before ingest

#6 Align Research Data Services

The University Libraries will unify research data staff and services around the user for an optimized user experience.

Marketing

- Create a landing page of all data services linked from libraries website
- Develop a marketing and communications plan

Staff Training

- Create a comprehensive training program for interested library staff

Organizational Structures

- Reassess the current organizational structures for providing data services in ways that better align services around the end-user
- Leverage/coordinate ticketing systems and increase communication across units
- Establish stronger roles and rewards for library staff that better reflect our matrix organization

Referral and Collaboration with Campus Units

- Test/pilot new models of cross-collaboration
- Develop a strong, pro-active referral network of existing data services on campus, e.g., the University Storage Council (USC) and the Research Data Management informal community of practice (RDM iCoP)

Background

The University of Minnesota Libraries has historically provided strong research data services (RDS) to campus users, providing significant services (see Fig 1) in the early planning stages (e.g., DMP consultations, discovery/purchase of data sets for reuse) and later stages (e.g., DRUM repository and curation services for data), and our [well-recognized Data Management web site](#) has reached thousands of users worldwide.

Figure 1: The University Libraries currently provide a number of Research Data Services

Finding/Using Datasets	Access and discovery services to purchased datasets. Ex. Business market research data; GIS data.
Best Practices and Training in Data Management Topics	On-demand workshops taught throughout the year. Ex. Bi-annual “Data Management Bootcamps” for graduate students; Course-embedded workshops for 20+ research methods classes; Managing Your Data website; Reproducibility portal .
Data Management Plan (DMP) Services	Provide templates and timely review of draft DMPs, a required component of many external funding programs. Maintain up-to-date list of funder public access policies .
Data Repository and Curation Services	Dataset archiving and hosting in the Data Repository for the University of Minnesota (DRUM) , services include long-term preservation and public access (DOIs). Data curation services available prior to public release (required via DRUM and on request for external repositories).

Consultations/Tools for Managing Data (DASH)	Consultations around emerging digital tools and methodologies for creating, managing, and publishing data. Ex. Building a digital library to showcase research in a sub-discipline; support for creating a content-rich website; hosting software carpentry workshops; consulting with basic scripting tools (R, Python).
More Data Related Services Offered in the Library....	Specialty Software and Scanning Equipment (e.g, SMART Learning Commons) Productivity Tools for Managing Citation Data (e.g., Mendeley, Zotero) Media Production Services (e.g, Video & Image Data) Digital Library Services (e.g, archival digitization services) Tech-Enabled Spaces in the Library (e.g., Makerspaces, Data Visualization Center)

Today, the growing volume of researcher-produced data on campus, combined with increased competition for funding, demands more aggressive data services, education, and intervention by the University Libraries and strategic partners on campus and beyond. A library staff retreat held on July 10, 2017 brought together library staff from across units to discuss the future of RDS in the libraries. We brainstormed trends that affect our library services (national, research, academia, government, private sector) and concluded that there were a myriad of RDS activities that the library could focus efforts in the near-term future (see appendix A)

Following the staff retreat, the University Libraries charged a Research Data Services Strategic Planning Task Force (see appendix B) to develop a strategic plan for Libraries' research data services to be implemented over the next 1-3 years. The plan should identify and justify which research data services are near-term priority for the Libraries, roadmap how to move forward effectively providing those research data services (staffing, structure, resources, process flow etc.), and acknowledge research data service areas that are important but best managed by other units and campus services. The task force is expected to engage with non-library stakeholders at key points in their process. The scope and specific tasks are discussed in more detail below.

The RDS Strategic Planning Task Force met weekly from November 2017 through February 2018 and performed tightly focused research, both internally and with external stakeholders. This report presents our methods to analyze and prioritize current and potential research data services across the entire data lifecycle and presents our strategic roadmap (see Table 6) for RDS that will best meet researcher needs and can make the greatest use of library expertise and resources for the near-term future.

Methodology

The work of the Task Force was broken into three stages: (1) Information collection and analysis about data needs and current service implementation with significant input from library staff; (2) engagement with influential campus stakeholders external to the library on

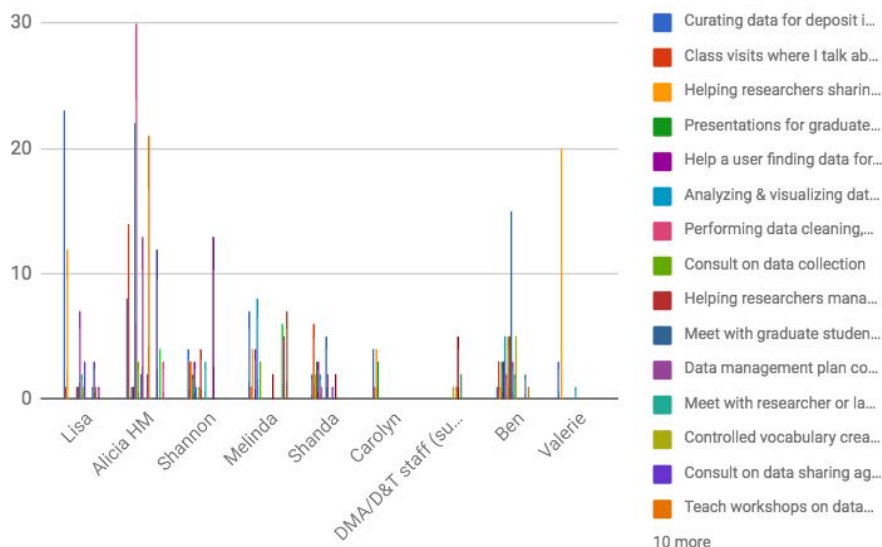
potential RDS offerings going forward; and (3) prioritization and presentation of recommendations to library leadership. The outcome is a vetted, fully-actionable strategic plan for Libraries' Research Data Services (see [Results Section](#)).

Stage 1: Information Collection and Analysis

Task Force members gathered and analyzed a significant amount of observational evidence (qualitative and quantitative data) about researcher needs related to data services. Observational evidence included:

1. **Direct Requests:** 108 examples of data-related consultations and/or reference requests from 2017. These self-reported examples were provided by library liaison staff in the Research Data Services Team ([link to members](#)) and through interviews with select librarian staff (e.g., DASH, business reference, DMA, HSL liaisons, etc.). These reference questions ranged from data analysis support for text analysis or big data to data sharing and preservation support (see examples in sidebar). And a few staff reported the frequency of such requests and consultations over a one-year period (see Fig 2). Activities like consulting with researchers around organizing their files and sharing their data in DRUM were more common for library staff than data analysis or visualization consultations, which were occurring in CLA-LATIS.

Figure 2. Data-related reference and consultation activity frequency in 2017 as reported by 9 staff



SIDEBAR: Example Library Consultation Requests in 2017

"I am going to get some money from NSF to put data buoys in Elk Lake and Itasca up at IBSL and another one in Cedar Bog Lake at Cedar Creek. These systems will be generating lots of data--several instruments will record wind speed, humidity, light levels, dissolved oxygen, pH and about a million other variables probably hourly for most of the ice-free season and probably some of it under ice. So, lots of data. I know the library assists folks with data management design and I think it would be good for me and my tech to talk to some folks about such things."

- Faculty member in Ecology, Evolution, and Biology

"I'm publishing some research. The journal requires that I archive the data. I think you can do this. I need to know how and, in particular, I must submit an accession number (or DOI) to the journal, or else they won't publish!"

- Faculty member in Kinesiology

"Is there depository for storing data and other materials from projects (research and otherwise) at the library that is NOT publicly accessible? The researcher for whom I work has talked about putting things in "deep storage," but I didn't think MLAC took physical materials from old research projects. Is there someone I could talk with about these questions?"

- Faculty member in Professional Programs.

2. **Staff Observations:** 84 data-related activities were brainstormed and mapped across the lifecycle at a two-hour meeting held November 15, 2017 for members of the RDS Team and invited library staff that regularly work with data-related issues. The outputs of this meeting fell into three categories (see Fig 3)
- Things we SAY we do (e.g., advertised services found on library web pages)
 - Things we ACTUALLY do (e.g. happening at the individual staff level but not necessarily advertised)
 - Things we should/could do (e.g., potential areas of unmet need).

Figure 3. Excerpt of Library Staff Brainstorming Results from RDS Team+ Meeting

Things we SAY we do	Things we ACTUALLY do	Things we should/could do
TEACH Education and Training in Data Management CONSULT Writing Data Management Plans Metadata Standards (not many requests for this) Scripting (R, Python) Selecting/Using Productivity Tools for Managing Citation Data PROVIDE Help find Existing Data (using library resources) Create/Use GIS data Create Video & Image Data Scanning & Digitization Services (for data analysis) Develop metadata schemas and standards (one in DMA) Data Repository and Curation Services (DRUM) Grant Partner Support	TEACH Teach in the classroom/workshops (e.g., Methods class, targeted course-embedded) (bi-annual) Data Management Bootcamp for Graduate students Training in data visualization tools (e.g., story maps) Training in text analysis tools Library staff education CONSULT Data Management Plan Reviews Federal DMP requirements and data sharing guidelines (OSTP memo) Data sharing broadly Overall research workflows/ Embedded data management (consulting with a lab) Selecting/using electronic lab notebooks (ELNs) File management as relates to workflow Managing paper/analog data Sensitive data storage - refer requests for storing human subjects data (not sharing) Personal information management and citation data De-identification techniques Data use agreements Language for IRB participant agreements Data licensing, how to select one Copyright, ethics & data External data repository selection (but not many requests for this) Non-affiliate requests for data curation (e.g. help with documenting and archiving data geospatial data) Metadata, controlled vocab for discovery Sharing Geospatial data (USpatial) Data citation challenges WRT publications (e.g. workflow of release) PROVIDE Find/Use bibliographic data Create/Manage Systematic review data Zooniverse data support	TEACH Teach reproducible workflows/techniques in labs/groups General understand of reproducibility and link with data management CONSULT Consult on end of project, project close-out needs Reproducibility consultations/referrals Support data/code citation PROVIDE Host OSF instance Host ELN tools/offer support (e.g. GitHub) Research reproducibility services Provide archival "dark" storage Data visualizations services Controlled vocabulary hosting Peer review for data (blind, double blind) Restricted access to data, beyond embargo period Catalog DRUM records in library discovery service layers Track impact measures of data sharing in DRUM (citations, etc.) Open data incentives tied to DRUM (e.g., badges) Hosting/storage of vended data (portico?) Data catalog/registry and access to purchased data Data integration (clearing, merging, making two datasets interoperable) Reproducible workflows for software and code STAFF NEEDS Expertise for consulting with scripting (R, Python) outside of CLA Text mining expertise Expertise in R to support bus/financial data Expertise in data visualization Better knowledge of licensing OTHER Rebrand data management stuff around reproducibility

3. **Meta-analysis of Campus Need:** Acting on the 2017 Library Staff RDS Retreat’s #1 crowdsourced recommendation to “Examine all collected data that discusses campus needs to find themes (surveys, environmental scan, literature, focus groups) to ensure services match need.” The team analyzed 24 prior reports and raw survey data that captured the research data needs of University of Minnesota researchers. Qualitative survey feedback was also summarized by team members when available. These included:

- | | |
|---|---|
| a. Ithaka Agriculture U of M Faculty Interviews, 2016 | j. U of M Libraries Research Services Coordinators Report (2017) |
| b. U of M Academic Health Center Faculty Survey, 2014 | k. U of M Libraries' Data Management/Curation Initiative (DMCI) Report, 2014 |
| c. U of M Biosciences Faculty Survey, 2016 | l. U of M Libraries' Electronic Records Task Force Report, 2017 |
| d. U of M Biosciences Grad Student Survey, 2016 | m. U of M Medical School Faculty Interviews, 2014-15 |
| e. U of M College of Forestry, Agriculture, and Natural Resources (CFAN) Faculty Survey, 2015 | n. U of M Ithaka Public Health Faculty Interviews, 2016 |
| f. U of M College of Liberal Arts (CLA) Faculty Interviews, 2013-14 | o. U of M Research Cyberinfrastructure Alliance Environmental Scan, 2009 |
| g. U of M College of Science and Engineering (CSE) Faculty Survey, 2014 | p. Kooper et al. (2017). “Research Data Services Maturity in Academic Libraries.” <i>Curating Research Data</i> . |
| h. U of M Digital Arts, Sciences, and Humanities (DASH) Library Liaison Survey, 2017 | q. U of M Research Data Services Retreat for Library Staff, 2017 |
| i. U of M Libraries Research Data Services Team discussion, 2017 | r. U of M Secure Storage Service Development Group, 2015 |
| | s. UCCS Storage Restructuring and Reorganization Committees (four reports), 2017 |

We also analyzed Data Management Bootcamp feedback. The bi-annual bootcamp has been delivered five times to mostly graduate students and post-docs since 2015. In total we analyzed 781 responses to the pre-workshop question “What do you want to get out of this workshop.”

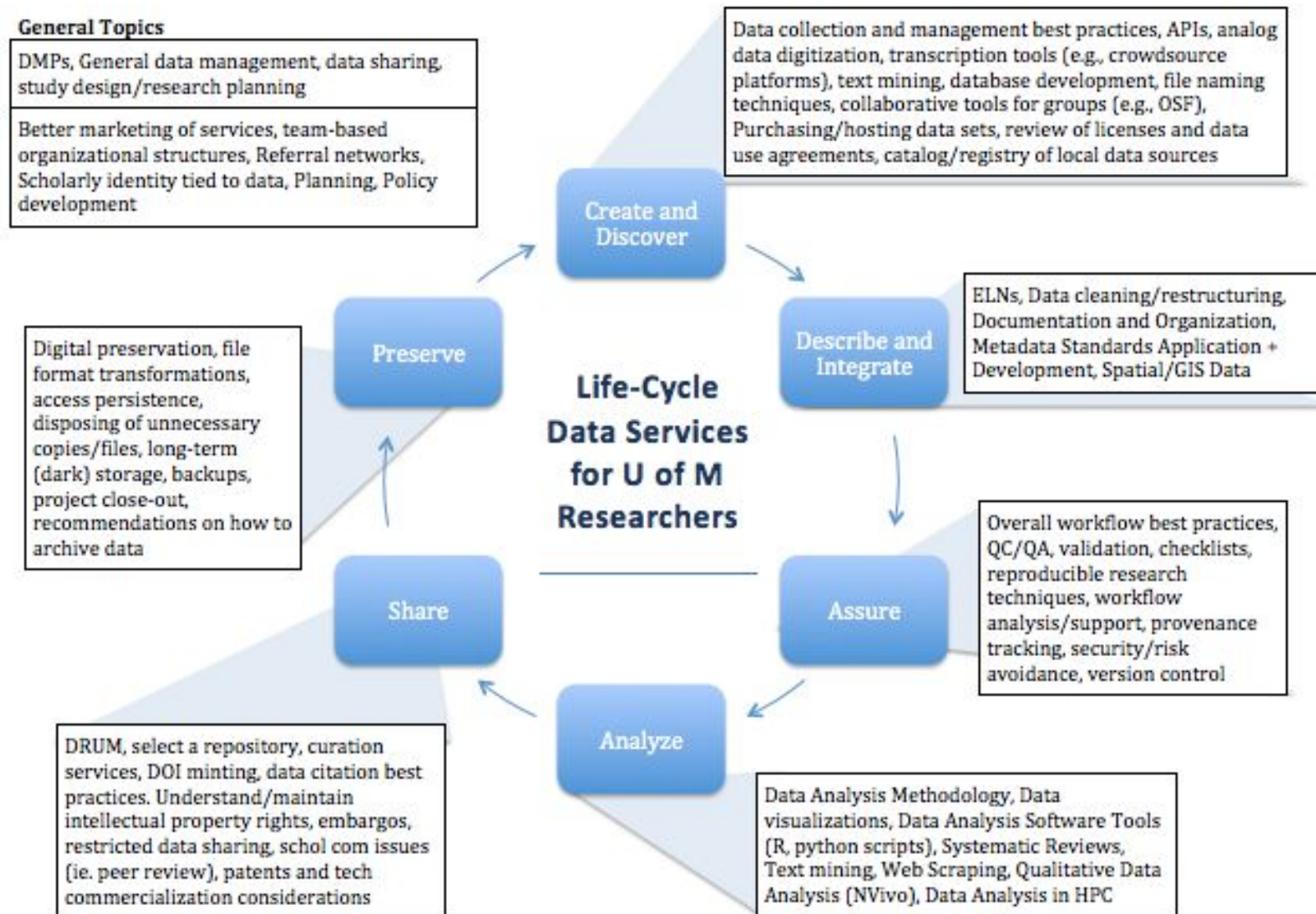
Stage 2: Stakeholder engagement around preliminary analysis

Next, task force members held interviews with identified campus stakeholders external to the library. We presented them with a visual of our ideas for research data support and services (see Fig 4) based on our research in Stage 1 and asked them:

1. Reviewing the summary of user-needs related to data services on campus, what needs resonates with your own experiences?
What do you see missing?
2. What additional areas would you like to see that would be relevant to you and your colleagues/stakeholders?
3. Do you see any additional partnership roles where the library should be collaborating with your unit or other units on campus?
4. What data services currently provided by the Libraries are you aware of? (help us to know where to focus marketing)
5. Beyond what you believe that the Libraries may or may not be able to do, what about data interests you; what trends excite you; what direction would you like to go with your data?
6. What other offices on campus would you expect to offer data services?
7. Describe what institutional support for data management would look like for you.

Overall, the interviews pointed to a glaring need for better marketing of Libraries data services and possible rebranding. The interviewees offered several "outside the box" solutions to expanding data services, as well as positively encouraged the Libraries to be more interdisciplinary and coordinate with other departments on campus. There was also strong interest in a resource that would help campus units navigate the wide array of services already in place, as well as any developing services. People were often unaware of resources, or unsure where to begin.

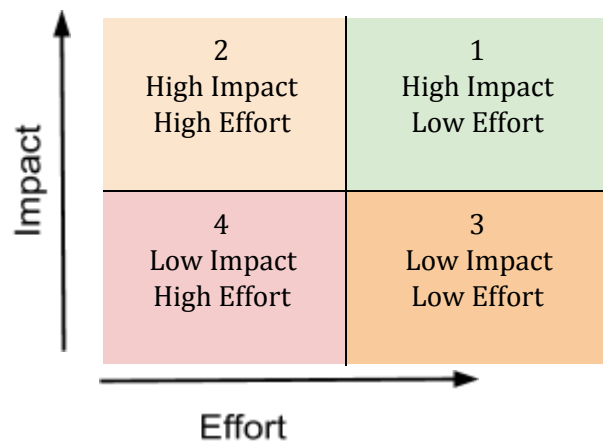
Figure 4: Handout of 300+ questions/consultations in 2017 we find that U of M Researchers have many needs related to data



Stage 3: Finalize data services strategic plan

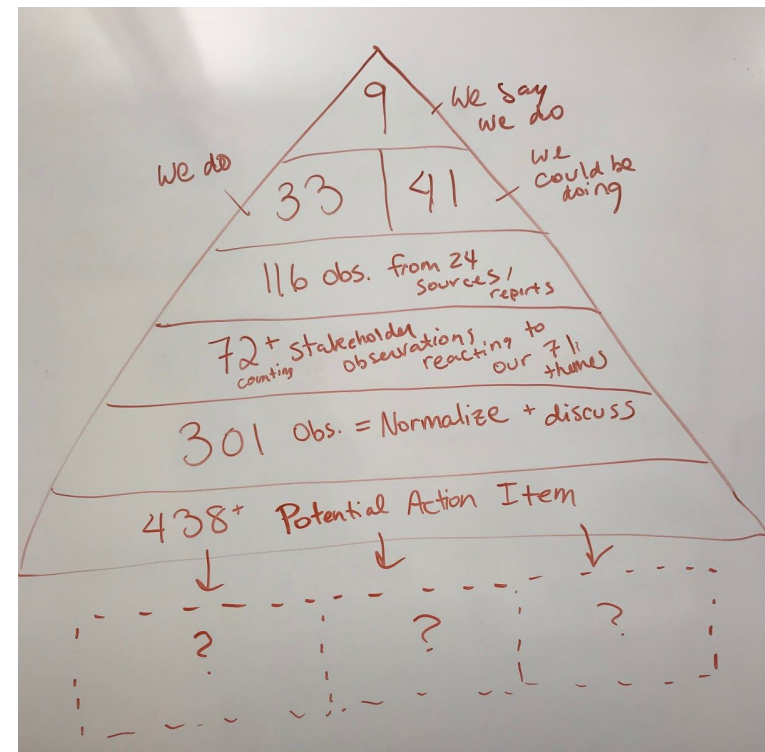
In January/February the team held meetings on nearly a weekly basis for 2-4 hours each to analyze our findings. The discussions were rich and informative. Taking the observational data collected in Stages 1 and 2, we normalized the observations into “themes” such as Electronic Lab Notebooks or Text Mining. All team members reviewed the themes and in our meetings we expanded or collapsed thematic categories to arrive at a total of 71 themes. Next, each member was assigned ~10 themes to brainstorm action items for each theme based on the detailed data collected to present and discuss with the group. **This resulted in 438 potential action items for the Libraries RDS roadmap** (see all potential action items in Appendix A).

To help prioritize the themes (and our discussions), we coded each theme and action item by perceived effort (cost to implement and/or staff time to implement) and perceived impact (reaching many people vs reaching only a few people). The resulting priority categories fell into a two-dimensional grid.



Next, each team member pulled their top five themes (out of the whole 71) that they felt were most crucial to the strategic plan and coded each action item listed for those themes. This double method of coding allowed us to compare prioritization ranking among team members which we found to be well-aligned.

Figure 5: Data collected by Task Force



Results

Out of the 71 themes, 18 rose to the top of our priority list through ranking and discussion. We also re-reviewed our discussions with our stakeholders to identify any missing themes in our top 18. The resulting top themes were categorized into six goals. We closely examined the action items and edited them according to feasibility, importance, and the discussions to date. Finally, we brought our draft results to the sponsors on February 26, 2018 for feedback. We added more detail to the presentation of the recommendations and submitted the final version for cabinet review on March 13, 2018. The results is resented in Figure 6 as our strategic plan.

Fig 6: University of Minnesota Libraries Strategic Plan 2019-2020

**Note: Evidence of need highlighted in this table represent a small subset of examples from our 400+ observations ([see methodology section below](#)). They do not have a direct 1-to-1 correlation with the ideas for action column.*

#1 Enhance Support Across the Research Workflow The University Libraries will enhance existing research data services offered across the entire research workflow.		
Theme	Ideas for Action (with Priority Ranking)	Evidence of Need*
Data Management	<ol style="list-style-type: none"> 1. High Impact/High Effort Establish core set of data librarians managers who can provide "immersed" data management support during the research lifecycle. This can range from one-shot presentations in research labs to going into researchers labs before a project, looking at their workflow, talking through file naming / organization, thinking about end goals of sharing (if for a grant). This could be branded as "Rent a data manager" or similar service. 2. High Impact/Low Effort Encourage liaisons to incorporate data management into interactions with students and their users by reframing data management into more accessible skills, such as "5 things every librarian can use in a 	<p>Need help managing data over the course of a project, restricted access when necessary - Secure Storage Service Development Group, 2015</p> <p>Lack of awareness around data management, need help actually doing it - Storage Restructuring and Reorganization committees, 2017</p> <p>Consult request: "I have recently received funding from the National Science Foundation for a project that includes multiple types of data. I am hoping to get some advice about setting up a solid data management system before the data start coming in from our first project activities. I hope to meet with someone from your office sometime in July." - Liaison Consults 2017</p>

	<p>consultation” (e.g., file naming/organization, mindful data storage, citation management as data management, etc.)</p> <p>3. High Impact/Low Effort Develop guides or workshops for best practices for managing archival data (analog, physical specimen, etc.) to better understand need/demand for this type of support.</p>	<p>Most frequent scenario for me: during one-on-one student consultation on literature searching, they'll ask about data management best practices (e.g., citation managers, file naming, storage) - Liaison Consults 2017</p>
Study Design / Research Planning	<p>4. High Impact/Low Effort Develop guidelines to help researchers submit IRB protocols that allow for responsible sharing and archiving of data. Work with IRB (HRPP) to add modules to CITI training to support best practices around data sharing and data security.</p> <p>5. High Impact/Low Effort Coordinate with grants coordinators to advertise available grant opportunities and grant preparation/writing services across campus. Align grant funding workshops with broader research planning topics such as DMPs.</p> <p>6. High Impact/Low Effort Familiarize relevant library staff (RDS staff, liaisons to departments doing human subjects research) with IRB processes to facilitate consultation on IRB submissions, coordinating with other offices as applicable (IT, collegiate research offices).</p>	<p>Advise the Libraries to run pilots, Look for experts in existing departments - SPTF Stakeholder interviews, 2017</p> <p>Library services overlap with theirs (AHC specific) - SPTF Stakeholder interviews, 2017</p> <p>We do this: Assistance with grant funding - Libraries Research Services Report (2017)</p> <p>Libraries role identified to helping researchers design their research, from broad to specific (surveys) and help with navigating IRB process. - Research Cyberinfrastructure Alliance Enviro Scan, 2009</p>
Data Management Plans	<p>7. High Impact/High Effort Work with Sponsored Projects Administration to regularly be alerted of funded DMPs from researchers (also provide to relevant college office). Review funded UMN DMPs and provide timely feedback.</p> <ul style="list-style-type: none"> Establish a timeline to systematically follow-up with those who included DRUM in their plan. <p>8. High Impact/Low Effort Make an “actionable DMP” that includes a timeline for updating DMP, training staff in best practices, consulting with a data librarian, and sharing data via appropriate repositories (e.g., DRUM).</p> <p>9. High Impact/High Effort Work with Grad school/GSSP to</p>	<p>Need to see DMPs to know what the responsibilities are for data, esp when people leave and currently do not - SPTF Stakeholder interviews, 2017</p> <p>Strengthen relationship with OVPR - a lot of information about DMPs go through SPA and never get seen by people who need to know - SPTF Stakeholder interviews, 2017</p> <p>Send researchers to library for this - templates useful - SPTF Stakeholder interviews, 2017</p>

	create DMP requirements for all students writing a dissertation or thesis. Consult with a librarian as part of regular thesis planning.	
Reproducible Research	<p>10. High Impact/Low Effort Rebrand data management services to better reflect growing interest/awareness/demand around reproducibility (e.g., promote DRUM as an "open data" tool, branded instance of OSF).</p> <p>11. High Impact/High Effort Pilot specific services to support reproducibility (e.g., code review/debugging services) in conjunction with other campus groups rather than going alone (feedback from stakeholders), for example the Computer Science Department.</p> <p>12. High Impact/Low Effort Identify example workflows that support research reproducibility best practice and promote via websites (e.g., Land of 10,000 workflows, Reproducibility Portal, ELNs), and provide examples for workshops/trainings.</p>	<p>There is an attack on science - we must be active vs reactive - SPTF Stakeholder interviews, 2017</p> <p>Library should get involved in reproducibility and pre-trial registration - SPTF Stakeholder interviews, 2017</p> <p>Libraries should take lead on reproducibility - SPTF Stakeholder interviews, 2017</p> <p>Facilities need support in this - UMII does some, but not something the library can take on so easily. Should focus on stuff U can standardize, like telling researchers to use GitHub - SPTF Stakeholder interviews, 2017</p>
#2 Provide Support for New/Emerging Needs The University Libraries will recognize and respond to new and emerging data-related needs that our campus researchers face.		
Theme	Objectives	Evidence of Need
Data Use Agreements (DUAs)	<p>13. High Impact/Low Effort Partner with relevant campus offices (e.g., Unfunded Research Agreements) to provide consultation for:</p> <ul style="list-style-type: none"> ○ DUAs for requesting existing data from an external source ○ Writing DUAs for data the researcher/University produces to be shared to others 	<p>Almost all researchers have external partners for data collection - SPTF Stakeholder interviews, 2017</p> <p>OTC, OGC see DUAs as a need - SPTF Stakeholder interviews, 2017</p> <p>Concerns around security and trust with data sharing; Controlling access to data that are shared -</p>

	<p>14. High Impact/Low Effort Implement custom DUAs in DRUM to allow more flexible sharing, such as custom terms of use or credential login to access data files.</p> <p>15. High Impact/High Effort Explore potential levels of support for sensitive data in DRUM by allowing customizable data use agreements that users must agree to or potentially register to use).</p> <p>16. Low Impact/High Effort Permanent embargos for data that requires approval from a centralized manager (in libraries?).</p>	<p>Agriculture Faculty Survey, 2016</p> <p>Need access granted case-by-case, restricted data sharing - AHC Faculty Survey, 2014</p> <p>Restricted data sharing - want control, submit proposal before access is granted - CLA Faculty Survey, 2013</p> <p>"My question is about multi-site sharing of human subjects data including videos. Is Google Drive considered secure enough to protect identifiable data, and if so can it be used across multiple institutions?" - Liaison Consults 2017</p>
Data Visualization	<p>17. Low Impact/High Effort Provide support for creating public-facing data visualizations, including design consulting, web-hosting, etc. Pilot new service models with other campus groups rather than going alone (feedback from stakeholders), for example the College of Design and Computer Science both work on data visualization and reproducibility issues, respectively.</p> <p>18. High Impact/Low Effort Provide access to a range of licensed data visualization tools.</p> <p>19. High Impact/High Effort Develop expertise in how to select and use data visualization tools; partner with others on campus that host visualization tools as appropriate.</p>	<p>Overlap with college - CEHD has an ed tech visualization office, people in CAREI also interested in providing this - SPTF Stakeholder interviews, 2017</p> <p>Scientists don't know how to tell their story. Scientists need a story, impact, and communication - however their focus is their methodology and publication. - SPTF Stakeholder interviews, 2017</p> <p>"Need more support for 'preparing data for presentation,' I mean those tasks which I cannot do myself. This means high-end graphics or animations which many of my competitors have in their presentations." - CSE Faculty Survey, 2014</p>
#3 Build on Successful Education/Training The University Libraries will scale successful training programs in research data management to reach a broader audience.		
Theme	Objectives	Evidence of Need

General	<p>20. High Impact/Low Effort Better market data management education opportunities in the library (see Marketing) as well as provide a platform to promote events across campus.</p> <p>21. High Impact/Low Effort Collaborate with campus programs, such as the Undergraduate Research Opportunities Program (UROP) or the Honors program, that present opportunities to introduce undergraduate students to data management best practices.</p> <p>22. High Impact/High Effort Incorporate data management planning into discussions/trainings for theses and dissertations to better prepare students with the skills and experience needed in the e-research environment.</p> <p>23. Low Impact/High Effort Provide immersive professional development opportunities for grad students through data curation fellowship or internship, introducing data curation and other "alternative academic (alt ac)" career opportunities.</p>	<p>Need for professional development training clearinghouse of campus opportunities for training. - SPTF Stakeholder interviews, 2017</p> <p>Add DM training module to CITI training - only reaches new folks - if we tacked our training onto HRPP though there would be a refresher - SPTF Stakeholder interviews, 2017</p> <p>Need to focus more on undergraduate education for RDM - SPTF Stakeholder interviews, 2017</p> <p>An important training need for grad students. Strong need to scale this to reach more grad students. Continued interest in CFANS-specific DM bootcamp - SPTF Stakeholder interviews, 2017</p>
One-shot Workshops	<p>24. High Impact/High Effort Target data management education to research support staff in specific labs (facilities) or research groups.</p> <p>25. High Impact/Low Effort Expand the graduate student-focused "Data Management Bootcamp" training by moving it into Canvas or Destiny (for better registration/lesson planning) to better scale and offer certificates.</p> <p>26. High Impact/High Effort Engage with campus deans and related stakeholders on education issues surrounding data management. Develop partnerships within the colleges to promote this training.</p>	<p>CEHD lags behind in Data Management; big need for outreach/education - SPTF Stakeholder interviews, 2017</p> <p>Target junior faculty. Target facilities vs PIs. Need DMPs and curation training for facilities - SPTF Stakeholder interviews, 2017</p> <p>Need more workshops on R, Python, Github and reproducibility - SPTF Stakeholder interviews, 2017</p> <p>Maybe have a certificate program for this? - SPTF Stakeholder interviews, 2017</p>
Course Embedded	<p>27. High Impact/Low Effort Better advertise "one-shot" presentations that can be delivered in classrooms or</p>	<p>Data management training is used a lot. Grad students are clamoring for it. His class filled up immediately (data</p>

Workshops	<p>tailored for specific research groups/labs. Coordinate and expand staff who are able to give them.</p> <p>28. High Impact/Low Effort Give internal library staff presentation on how we've worked with methods courses in the past to get other liaisons involved as R&L goal for liaisons to reach out in a similar way to their departments.</p>	<p>management and related topics). Need more training and need to scale training. - SPTF Stakeholder interviews, 2017</p>
#4 Facilitate Data Access The University Libraries are exceptional providers of content in all formats, including data, and will enhance users' access to our growing collections of digital research data.		
Theme	Objectives	Evidence of Need
Data Collection Development and Management	<p>29. High Impact/High Effort Develop a sustainable workflow to provide delivery of data sources that the library purchases/licenses on behalf of the university/researchers and to reduce potential for redundant dataset purchasing across campus (e.g, LIDAR data).</p> <p>30. High Impact/Low Effort Gather information from RDS members and others in Libraries (Electronic Resources Management (ERM)) about candidate data sets (licensed, purchased, or free) for potential Libraries hosting; create a list that provides evidence for need.</p> <p>31. High Impact/Low Effort Meet with D&T and C&C folks to talk about methods and technologies (e.g., Portico, Globus) for hosting data sets for sharing. Reference procedures for providing access to licensed data to potentially provide more controlled access for researcher data.</p>	<p>Access to data sources that the library purchases/licenses on behalf of the university/researchers - Research Cyberinfrastructure Alliance Enviro Scan, 2009</p> <p>Need access to data sources that the library purchases/licenses on behalf of the university/researchers - Libraries' Data Management/Curation Initiative Report, 2014</p> <p>Providing access to purchased electronic library materials inaccessible to users (value for 2015-2016 fiscal year of nearly \$41,000) - Libraries' Electronic Records Task Force Report, 2017</p> <p>Large dataset (Google Drive) purchased and delivered as a bunch of files that we make available through the Bus and data lab (wilson collaboration studio) and cataloged (Historical company files) (\$60,000) - Liaison Consults 2017</p>

Discovery and Access to Data	<p>32. High Impact/Low Effort Create single entry website/landing page that unites the various data sources - map, government, text data, SCOPUS, campus data sets (e.g., data registry, etc).</p> <p>33. High Impact/Low Effort Advertise more aggressively that the libraries can help researchers find and ethically use data sets in their research.</p> <p>34. High Impact/High Effort Evaluate workflows for indexing DRUM datasets in library discovery services (either MNCAT or Primo), in consultation with D&T and D2A (Discovery to Access) Steering Committee to help set a foundation for the "data catalog".</p> <p>35. High Impact/High Effort Create a "registry" or "data catalog" of data sets which are available to researchers (locally hosted/published and licensed/purchased) [within existing discovery/access framework/workflows, not as separate database, etc.]</p> <p>a. Could just be a resource type within Alma/Primo then using the API / RSS push out to a separate interface (if desired).</p>	<p>Lots of university data that isn't being used well - accessibility, tracking - SPTF Stakeholder interviews, 2017</p> <p>Want a system to shared vetted/purchased data across departments...an Experts platform for research data services across campus - SPTF Stakeholder interviews, 2017</p> <p>Seeing datasets in the context of business information (typically want a consultation, contact via email, team projects) - Liaison Consults 2017</p>
<p>#5 Provide Services for Long-term Data Stewardship</p> <p>The University Libraries will ensure accessibility and preservation of research data through curation, metadata, repositories, and other access and retrieval mechanisms to meet federal, state, sponsor, and University requirements. (per U Policy)</p>		
Theme	Objectives	Evidence of Need
Project Close Out	<p>36. High Impact/Low Effort Offer specific guidance and consultation around project close-out, to help reduce data loss. Could offer "rent a data manager" style services to help clean-up project for sharing and archiving.</p>	<p>Need to make sure there are continuing exchanges between various Units on campus, Storage Council is part of this - SPTF Stakeholder interviews, 2017</p> <p>Losing data due to ineffective methods Public Health</p>

	<p>Coordinate with storage council and IT units who are archiving data.</p> <p>37. High Impact/Low Effort Directed advertisement/notification to graduating students about how to manage their data when they leave.</p> <p>38. High Impact/Low Effort Update thesis writing guidance with instructions on how to deposit data to UDC as supplemental file for thesis record or as separate DRUM record.</p>	<p>Faculty Interview, 2016</p> <p>Data are currently stored across many systems and hard to keep track of - Med School Faculty Interviews, 2014-15</p> <p>Consult on end of project, project close-out needs - RDS+ Listening Session, 2017</p> <p>Post-project help for long-term storage, files sit on server - CLA Faculty Survey, 2013</p>
Data Preservation	<p>39. High Impact/High Effort Groups on campus (e.g., Minnesota Supercomputing Institute, St Anthony Falls Lab) are providing archival data storage. Partner with these groups to prevent potentially costly maintenance of "data dumps" that even the author may not understand in 2-5 years.</p> <p>40. High Impact/Low Effort Raise awareness of the need for preservation of valuable data (i.e., unique, institutional assets) that goes beyond storage and backup (e.g. Rosetta).</p> <p>41. High Impact/Low Effort Create procedures (attached to RDM policy) preserving data that are informed by the UMN Libraries Digital Preservation Framework.</p> <p>42. High Impact/High Effort Work with campus stakeholders to implement/integrate archiving and preservation features into existing data storage systems to better support the research process.</p>	<p>Lack in the areas of long term storage and preservation - SPTF Stakeholder interviews, 2017</p> <p>Need long term data preservation assistance - AHC Faculty Survey, 2014</p> <p>Qualitative data, dark archives, large video files - how to preserve? - CLA Faculty Interviews, 2013-14</p> <p>Want guidelines and procedures for moving inactive data to less expensive storage - Storage Restructuring and Reorganization Committees, 2017</p> <p>Long term archiving of data is very important, especially with an eye to complementing the more computational mission of MSI. DRUM is a good start on this regard. - Biosciences Faculty Survey, 2016</p> <p>Most wanted more help preserving their data and storing/backing up their data. They also are very savvy, want to know the protocols, backup procedures, and migration schedules to understand how well their data is being preserved. CSE data can be useful for centuries to come, notes one fac in astronomy. - CSE Faculty Survey, 2014</p>

Data Repository and Curation	<p>43. High Impact/High Effort Partner on building infrastructure internally that supports the data curation process for data "archived" in other university spaces (e.g., direct push to DRUM from MSI via Globus; Elevator to DRUM).</p> <p>44. High Impact/Low Effort Build out workflow to easily enable authors to submit data upon submission of manuscript to a journal that requires data sharing. "Under review" status for data that indicates it could change to reflect reviewer feedback on manuscript/analysis (prevents need for versioning). Provide mediated access to data and code that enables reviewers/editors to access the data confidentially (without the author needing to field the request), but is not publicly available until manuscript acceptance.</p> <p>45. High Impact/High Effort Create a staging platform for manipulating DRUM files before upload while capturing consent from authors.</p>	<p>How to deal with sharing and archiving big data - SPTF Stakeholder interviews, 2017</p> <p>DRUM is good for public data. If demand is growing for big data, should explore ways to connect Globus. This could be an updraft grant as a way to build this connection - SPTF Stakeholder interviews, 2017</p> <p>"I was wondering if you can have a portal to the data release archive of MSI. All of our sequencing data are permanently store at their data release archive, but in order to upload to DRUM, I need to download them to my disk, compressed them, and upload again, which seem to be a waste of resources." - Liaison Consults 2017</p> <p>"I am looking to upload files that accompany a manuscript that we're about to resubmit for review by a journal. We'd like the files to be uploaded before we resubmit (in a few days) so that the manuscript reviewers can access them." - Liaison Consults 2017</p>
#6 Align Research Data Services The University Libraries will unify research data staff and services around the user for an optimized user experience.		
Theme	Objectives	Evidence of Need
Marketing	<p>46. High Impact/Low Effort Create a landing page of all data services and include data discovery and creation services. Incorporate into a new "Find Data" page linked from home page of libraries website. (e.g., Consider using lib.umn.edu/data)</p> <p>47. High Impact/Low Effort Develop a marketing and communications plan in conjunction with University</p>	<p>Survey indicated low awareness of MSI, Libraries, and collegiate data services - CFANS Faculty Survey, 2015</p> <p>Lack of awareness of resources available - SPTF Stakeholder interviews, 2017</p> <p>Only data service aware of is finding data sets - SPTF Stakeholder interviews, 2017</p>

	Libraries Communications Office and other internal groups (e.g., Schol com, DASH) that reaches a broader audience.	<p>Change our image - what we list in the first part of lifecycle doesn't seem "library-like" - SPTF Stakeholder interviews, 2017</p> <p>Market to junior faculty, training grant scholars - SPTF Stakeholder interviews, 2017</p>
Organizational Structures	<p>48. High Impact/Low Effort Reassess the current organizational structures for providing data services in ways that better align services around the end-user rather than via individuals and separate library groups. These include the RDS Team/DRUM Curators subgroup and data-related library structures (Research Services Coordinators, Personal Information Management, Scholarly Communications Committee, etc.).</p> <p>49. High Impact/High Effort Develop workflows for referring requests that would benefit from expertise distributed across units in the library. Leverage/coordinate in-place ticketing systems and increase communication across units about work being done and expertise available in-house.</p> <p>50. High Impact/Low Effort Establish stronger roles and rewards for library staff that better reflect our matrix organization. For example, library staff interested or involved in research data service work should be rewarded in the annual review processes for collaborating on solutions that scale to our campus community rather than solely on individual contributions (e.g, papers, etc.).</p>	<p>Borrowed expertise model, bring staff across units to consult on project. Using tools to track commitment and time. - DASH Library Liaison Survey, 2017</p> <p>Many committees/groups have overlapping functions, but haven't had formal coordination in the past - Libraries Research Services Report (2017)</p> <p>Internal skills inventory - RDS+ Listening Session, 2017</p> <p>Formal partnerships within and outside, support from university administration - Research Data Services Maturity in Academic Libraries article, 2017</p>
Referral and Collaboration with Campus Units	<p>51. High Impact/High Effort Test/pilot new models of cross-collaboration, centralized work with focused topics.</p> <p>52. High Impact/High Effort Develop a strong, pro-active referral network of existing data services on campus and create</p>	<p>Need to be clearer who offers what. UCCS - data storage council good start in this direction. Need a 'Data management' referral network also. Maybe expand scope of this group? - SPTF Stakeholder interviews, 2017</p>

	<p>awareness of these cross-departmental services amongst services providers and as a result, library staff can more naturally and actively pull in other service providers during consultations.</p> <ul style="list-style-type: none"> a. Formal: Library and university representation on UCCS subcommittee "Data Storage Council" charged with addressing storage education recently formed to help address storage issues. b. Informal: Reboot the informal community of practice to allow service providers across the University come together to discuss Data Management issues. <p>53. High Impact/Low Effort Develop a marketing plan that is inclusive of cross-college collaborations on services in how they're promoted.</p>	<p>Coordinated effort to reduce gaps in research data services - waste of resources to have everything siloed. We should look to other universities for how they centralized services - SPTF Stakeholder interviews, 2017</p> <p>Lots of room for institutional collaboration - University Librarian as ambassador - SPTF Stakeholder interviews, 2017</p> <p>See potential for collaboration between libraries and many groups - SPTF Stakeholder interviews, 2017</p> <p>Need to make sure there are continuing exchanges between various Units on campus, Storage Council is part of this - SPTF Stakeholder interviews, 2017</p> <p>Don't have relationship with liaisons/libraries - SPTF Stakeholder interviews, 2017</p> <p>Many IT departments are now in more specialized research spaces - good to partner with them, help others transition - SPTF Stakeholder interviews, 2017</p>
Staff Training	<p>54. High Impact/High Effort Create a comprehensive training program for all interested staff with three levels of knowledge: introductory (best practices, awareness), intermediate (comfort with performing in person consultations), advanced (methods/data visualization techniques, curators, train-the-trainer).</p> <p>55. High Impact/High Effort Perform activities with library liaison staff to increase their better understanding of research design terminology (e.g., is the question about scientific methods vs workflows) in order to refer users to appropriate offices (e.g., other library staff or research and statistical offices).</p> <p>56. High Impact/Low Effort Familiarize library liaisons who</p>	<p>RSC trains liaison librarians in developing areas related to data services (internal training) - Libraries Research Services Report (2017)</p> <p>Including and engaging staff outside of liaisons in this work - Libraries Research Services Report (2017)</p>

	consult with researchers with service providers across campus through formal or informal mechanisms (see Align Data Services) to better help researchers navigate resources through their research (IT, college research offices, IRB, MSI, IRSA, etc).	
--	---	--

See full list of themes/action ideas that were discussed and considered in this process in Appendix A: Research Data Services 2018 “Themes”

Bibliography

Bakker, Caitlin. (2015). Interviews with 50 Med School Faculty at the University of Minnesota in 2015. Unpublished.

Bakker, Caitlin; Bishoff, Josh. (2016). Basic Sciences Assessment: In support of the Medical School & the College of Biological Sciences. Unpublished.

Calahan, Lisa; Kussmann, Carol; University of Minnesota Electronic Records Task Force. (2017). Electronic Records Task Force Phase 2 Final Report. Retrieved from the University of Minnesota Digital Conservancy, <http://hdl.handle.net/11299/189543>.

Farrell, Shannon L.; Kocher, Megan. (2016). Examining the Research Practices of Agricultural Scholars at the University of Minnesota - Twin Cities. Ithaka S+R Agriculture. Retrieved from the University of Minnesota Digital Conservancy, <http://hdl.handle.net/11299/183471>.

Hofelich Mohr, Alicia; Johnston, Lisa R; Lindsay, Thomas A. (2016). The Data Management Village: Collaboration Among Research Support Providers in the Large Academic Environment. *Databrarianship: The Academic Data Librarian in Theory and Practice* (editors Kristi Thompson and Lynda Kellam) Association of College and Research Libraries (ACRL). Retrieved from the University of Minnesota Digital Conservancy, <http://hdl.handle.net/11299/181127>.

- Hofelich Mohr, Alicia; Bishoff, Josh; Johnston, Lisa R; Braun, Steven; Storino, Christine; Bishoff, Carolyn. (2015). Data Management Needs Assessment - Surveys in CLA, AHC, CSE, and CFANS. Retrieved from the University of Minnesota Digital Conservancy, <http://hdl.handle.net/11299/174051>.
- Hofelich Mohr, Alicia; Bishoff, Josh; Johnston, Lisa R; Braun, Steven; Storino, Christine; Bishoff, Carolyn. When data is a dirty word. *ASIS&T Bulletin*, October/November 2015. <https://www.asist.org/publications/bulletin/oct-15/when-data-is-a-dirty-word/>
- Hunt, Shanda L; Bakker, Caitlin J. (2017). An Information Science Needs Assessment of School of Public Health Researchers at the University of Minnesota: Summary Report. Retrieved from the University of Minnesota Digital Conservancy, <http://hdl.handle.net/11299/189193>.
- Johnson, Layne M.; Butler, John; Johnston, Lisa R. (2012). Developing E-science and Research Services and Support at the University of Minnesota Health Sciences Libraries. *Journal of Library Administration*. Retrieved from the University of Minnesota Digital Conservancy, <http://hdl.handle.net/11299/159983>.
- Johnston, Lisa; Lafferty, Meghan; and Petsan, Beth (2012) "Training Researchers on Data Management: A Scalable, Cross-Disciplinary Approach," *Journal of eScience Librarianship*: Vol. 1: Iss. 2, Article 2. <http://doi.org/10.7191/jeslib.2012.1012>.
- Johnston, Lisa R. (2014). A Workflow Model for Curating Research Data in the University of Minnesota Libraries: Report from the 2013 Data Curation Pilot. Retrieved from the University of Minnesota Digital Conservancy, <http://hdl.handle.net/11299/162338>.
- Johnston, Lisa R; Bishoff, Carolyn. (2015). A Review of Data Management Plans (DMPs) from Successful National Science Foundation Grants from the University of Minnesota Twin Cities, 2011-2014. Retrieved from the University of Minnesota Digital Conservancy, <http://hdl.handle.net/11299/172728>.
- Kouper, I., Fear, K., Ishida, M., Kollen, C., & Williams, S. C. (2017). Research Data Services Maturity in Academic Libraries. In L. R. Johnston (Ed.), *Curating Research Data: Practical Strategies for Your Digital Repository* (Vol. 1, pp. 153-170). Association of College and Research Libraries.
- University of Minnesota. (2016). Box Secure Storage Needs Assessment. Unpublished.
- University of Minnesota. (2017). Storage Redesign and Restructure Committee (SRRC) Subcommittee 1: Service Analysis Storage Report. Unpublished.

University of Minnesota. (2017). Storage Redesign and Restructure Committee (SRRC) Subcommittee 2: Campus Needs Report. Unpublished.

University of Minnesota. (2017). Storage Redesign and Restructure Committee (SRRC) Subcommittee 3: Standards and Operating Procedures Report. Unpublished.

University of Minnesota. (2017). Storage Redesign and Restructure Committee (SRRC) Subcommittee 4: Education and Storage Champions Report. Unpublished.

University of Minnesota Libraries. (2015). The Supporting Documentation for Implementing the Data Repository for the University of Minnesota (DRUM): A Business Model, Functional Requirements, and Metadata Schema. Retrieved from the University of Minnesota Digital Conservancy, <http://hdl.handle.net/11299/171761>.

University of Minnesota Libraries. (2019). Research Data Services. <https://www.lib.umn.edu/datamanagement>.

University of Minnesota Research Cyberinfrastructure Alliance. (2009) Services Portfolio Matrix. Unpublished.

Appendix A: Research Data Services 2017 “Themes”

Themes around research data services and areas of support that the library may consider over the next 1-3 years.

1. Advocacy for Open Data
 - a. Scholarly communication consultation and advocacy for open data, data governance, research impact related to data.
 - b. Policy guidance for U of M on research data management (e.g., procedure development?)
 - c. Recommendations for licensing of open data/materials.
2. Analog Data (Digitization, Conversion)
 - a. Select one or two departments to conduct a “data audit” of departmental data (digital and analog). Consider the methodology of the Data Audit Framework, implemented at the University of Edinburgh.
 - b. Consider development of guidelines or service to convert/digitize data in paper form (within libraries collections or across campus, eg. DIY overhead scanners).
 - c. Research and develop guidelines to determine best practices for preserving analog data (e.g., keep analog but use archival standards, scan for preservation or convert to a machine readable format). Decision tree...since digitization is \$\$.
 - d. Work with Zooniverse to improve text transcription projects (advise on copyright issues)
3. APIs (connect to bib, text mining)
 - a. Build support for APIs for bibliographic data
 - i. This can include better front-end workflows for capturing requests from users and having staff run through the APIs
 - ii. Advertising available data by request
 - iii. Provide documentation and information on who can help with using APIs for library data (may not be available to make public) - includes connecting with electronic resources staff when data requested belongs to a database vendor.
 - b. Formal relationship building between library staff, so knowledge is shared (e.g., database API requests going to the right person to refer to vendors, etc)
 - c. Build up referral network based on experts who deal with certain types of data (SCOPUS, maps, gov, etc)
 - d. (inferred) Build APIs for other library data sources .
 - e. Document how to use existing APIs (e.g., DRUM, UDC, etc)

4. Backup: This theme belongs with data management themes around solutions for data backup
5. Bibliographic Data
 - a. Ensure that subject librarians know how to refer researchers seeking to create/use bibliographic data sets to DMA (ServiceNow request, or talk to Betsy.)
 - b. Make connections between DMA bibliographic data experts and citation manager tool experts so they can coordinate to meet researcher needs (DMA can help create data sets about books/journals, and they might get these requests in the course of their work).
6. Citation managers
 - a. Offer more in-depth instruction around personal information management in relation to citations
 - i. Reach out to citation managers group
 - b. Collaborate with graduate school to make citation management part of the thesis process
 - c. Bootcamp citation managers are popular
 - d. Create workflow for getting citations from Word/Excel into manager
 - e. Integrate the libraries better into the thesis workflow (citation managers, lit searches, DMPs)
7. Collaboration Tools
 - a. Enable UMN branded instance of OSF (shib login) - ambassador and config login
 - b. Use Data Bootcamp venue as a place to better educate future researchers on collaboration tools - lecture vs breakout session
 - c. Expand ELNs webpage to include collaboration tools and/or reproducible workflows
8. Course-embedded workshops: Better advertise "one-shot" presentations that can be given in classes or lab groups. Coordinate group of people who are able to give them. For example, customizable but mostly ready to go presentations on:
 - a. Data Management (RDS +)
 - b. Citation Managers (all Library)
 - c. Writing a Data Management Plan (RDS +)
 - d. Data Sharing in Sciences/Social Sciences, etc (RDS +)
 - e. Using the Library (all Library)
 - f. Research Projects (all Library)
9. Methods courses
 - a. Expand methods class outreach to more classes/colleges
 - b. Give instructions on how we've worked with methods courses in the past to get other liaisons involved as R&L goal. (CBS & AHC)

- c. Do a brownbag on what has been done already
- 10. Crowdsourced transcription
 - a. Create opportunities/ pipelines for researchers to use citizen science and transcription platforms for both textual and numerical data (Zooniverse @ U of M, fee-based services).
 - i. Case study about how projects get into zooniverse. What it does not do.
 - ii. Advise on when to use different approaches. Matrix
 - iii. Appoint a point person.
- 11. Citizen Science
 - a. Point person who can work in this area, with ties to data rescue and Zooniverse.
 - b. Hire a Citizen Science Librarian: Programming/tech knowledge, social science knowledge too
- 12. Curriculum Development
 - a. Identify ways to incorporate “data information literacy” techniques and skills into the curriculum and mapped to relevant courses and existing standards, such as the student learning outcomes.
 - i. Publish 12 DIL competencies for grad students on the website / handout as a talking point to use with faculty
 - b. Create an online DM educational tool for those not on campus
 - i. Work with an instructional designer to develop an online DM course.
 - ii. Package the RDM bootcamp presentations as “online course” - low barrier
- 13. Data Analysis High Performance Computing
 - a. Strengthen knowledge of MSI services, strengthen partnership for referrals to make sure people get what they needed at MSI.
 - i. Appointing someone in library to be “liaison” to MSI
 - ii. Develop user examples to give people an idea if MSI is right for them and have alternatives
 - b. Develop basic introductions to MSI services to help people know when to turn to MSI
 - i. Connected to Research workflow, could be tied to Data Management best practices for Big Data
- 14. Data Analysis Methods
 - a. Develop familiarity with new technologies/tools in data analysis, develop capacity for large scale consulting on tool selection, basic tool usage that could serve all disciplines.
 - i. Familiarity of tools vs familiarity of stats
 - b. Coordinate with partners on campus who provide data analysis/statistical support or advanced consulting in disciplinary uses of a tool/technology.
 - c. Identify partners or relevant courses that teach methodology

- d. Need more info: Review how people get support for data analysis in the colleges.

15. Data Analysis qualitative

- a. Advertise/adapt data management presentations and resources for qualitative research.
 - i. Consultations turn into teaching how to do qual research.
 - ii. Join interdisciplinary grad group for qual research
- b. Develop familiarity with technologies/tools in qualitative data analysis (e.g., NVivo), develop capacity for large scale consulting on tool selection, basic tool usage.
- c. Coordinate with partners for more specific disciplinary qualitative analysis or consultation - mixed methods group.

16. Data Analysis Software

- a. Gather information about college and University wide-licenses for data analysis software
 - i. Keep updated list to help researchers find what they have access to.
 - ii. Library needs to self-assess and ensure that special software is consistently offered across libraries' computers.
- b. Identify partners on campus with specific coding/analysis skills, take active role in connecting and coordinating these people across the U
 - i. Advertise and expand events such as pizza and programming that brings these partners together
 - ii. Data and Donuts that focuses on data analysis

17. Data Citation

- a. List clear guidelines for how to cite data and code in reference section of papers.
 - i. Templates or what select in zotero/mendeley.
- b. Add a data citation section to the librarian CV template.
- c. Increase availability of data citations in Experts@umn by populating from DRUM and UDC (currently there are 5).
- d. Research the best pipeline workflow for data sets submissions to DRUM that are associated with manuscripts and work out the best way to include the citation in the paper linking to the data. Share on website and DRUM site.

18. Data cleaning/restructuring

- a. Gather list of data cleaning tools (stats packages, coding tools, + others, such as OpenRefine) and resources for learning them
 - i. Expertise exists in DMA - might be able to suggest/consult?
 - ii. Share GitHub repositories of code
 - 1. Revisit policies and obtain clarity about sharing/open licensing practices
- b. Identify partners on campus with specific data restructuring skills in various tools, develop expertise in tools/disciplines where there are gaps.

19. Data Curation

- a. Expand expertise for data curation with a networked shared staffing model to collectively curate data among partner institutions (e.g., implement the Data Curation Network)
- b. Develop long-term solution to large datasets (curate or refer?) - Update our policy
 - i. Do we curate large datasets? How much checking is enough?
 - ii. What is a "large" dataset?
 - iii. At what point do we refer out and who do we refer to?
 - iv. Work with new archiving service at MSI (develop guidelines)
- c. Work with IT (OIT or collegiate or MSI) to identify inactive datasets that have been sitting on servers that would be good candidates for curation/archiving
- d. As a consultation, review "old" data and point out issues to inform the DM approach for new/future project data.
- e. Data collection development and management (where libraries store/make available licensed/purchased/curated data) [related to Discovery and Access to Data]
- f. Gather information from RDS members and others in Libraries (Electronic Resources Management (ERM)) about candidate data sets (licensed, purchased, or free) for potential Libraries hosting; create a list that provides evidence for need.
- g. Meet with D&T and C&C folks to talk about methods and technologies (e.g., Portico, Globus) for hosting data sets for sharing.
- h. Gather information about existing campus services that allow for data set transfer and sharing and explore potential for partnerships.
- i. Develop a sustainable workflow to provide access to data sources that the library purchases/licenses on behalf of the university/researchers.
- j. Reduce the potential for redundant dataset purchasing across campus (e.g, LIDAR data)
- k. Reuse procedures for providing access to licenses data to provide more controlled access for researcher data (DRUM; see restricted sharing).

20. Data management (see below too)

- a. Advertise data management help/consultation for archival data (analog data)
 - i. Develop guides or workshops for best practices for this, drawing from expertise in PIM
- b. Refresh liaisons on best data management practices (RSC?), integrate file naming/organization, mindful of storage into all consultations (citation management as data management, etc)
 - i. Rebrand away from data management to something like "5 things every librarian can use in a consultation"

- ii. Build competencies for DM that applies to all librarians - e.g., What are the research practices in your discipline? What type of data are generated?
 - c. Establish core set of data managers who can provide "immersed" data management support during the research lifecycle
 - i. Going into researchers labs before a project, looking at their workflow, talking through file naming/organization, thinking about end goals of sharing (if for a grant). "Rent a data manager"
 - ii. Maybe start at end of a project to help organize ("clean up" project for sharing and archiving), work backwards to devise better workflows? May be easier to target end of project folks (they know they need help)
21. Data Management Workflows (Embedded Librarian/Informationist)
- a. Partner with group/lab to develop end-to-end data management workflows and techniques (e.g., Center for Sustainable Polymers).
 - i. Idea of advertising services like this as "rent a data manager" - could also be lower bar lab workflow consultations (as done for several groups in SPH, Humphrey, Medicine, and College of Pharmacy)
 - b. Promote data management during the lifecycle (e.g., workflows, file organization, documentation, etc) - these are the most popular/impactful parts of our workshops and where we can offer consultation, but services don't currently reflect this well.
 - c. E.g., NIH Informationist grant program is a model.
 - i. Add-on grant to hire a librarian on a project (could model "rent a data manager" program on this)
 - ii. Lisa Federer paper on embedded librarianship case series
22. Data Preservation / "Dark" archival storage
- a. Raise awareness of the need for preservation of valuable data (i.e., unique, institutional assets) that goes beyond storage and backup (e.g, Rosetta).
 - i. Importance of curation through active stages of data lifecycle?
 - ii. Guidelines for disposal of data?
 - b. Provide dark archive/preservation services for data that are not public but should be maintained locally.
 - i. Tying into efforts in IT/MSI for archival storage? Encourage curation or at the very least thoughtful archival of data on theses systems to prevent potentially costly maintenance of "data dumps" that even the author may not understand in 2-5 years.
 - c. Implement/integrate archiving and preservation features into existing data storage systems to better support the research process.
 - d. Create guidelines for DIY preserving data (e.g., file formats for preservation, minimum level of metadata/description the author must create) that are informed by the UMN Digital Preservation Framework.

23. Data Registry [see: Discovery and Access to Data]

24. Data Repositories

- a. Research trustworthy data repository indicators and create a checklist for choosing the right repository for your data.
 - i. Define data repositories in not only their technical functions (dissemination, preservation) but also their mission, scope, longevity plans, etc..
- b. List a selection of data repositories for deposit that we recommend and why.
- c. Maintain a list of defunct or non-active data repository for edu purposes.
- d. Provide assistance to deposit data to external, disciplinary data repositories (i.e., ICPSR or GenBank) when appropriate.
 - i. Ambassador approach

25. Data rescue (Theme: Vulnerable or at-Risk Data in content, type, format)

- a. Data Rescue/Refuge/Libraries+ follow up - making connections with state and local libraries and archivist to understand current state of preservation and access to data produced by the state and local governments.
 - i. Civic switchboard (new Libraries+ project)
- b. Potential connections with other "at risk" initiatives - data on old formats, tied with old software, etc.
- c. Help faculty/staff/students save "at risk" data on old formats/software
- d. "Save Your Data Day" WITH DONUTS! - stations to help move old formats to modern storage options

26. Data Sharing

- a. Awareness of data sharing requirements from journals, agencies, etc.
 - i. Increase awareness of what public sharing means
- b. Develop set of online workflows and best practices to which we can point to when asked
 - i. Decision tree
 - ii. Sharing data for publication - create workflow for single blind peer review process of embargoed data in DRUM

27. Data sharing agreements (Consent forms)

- a. Update our IRB and consent form boilerplate language and templates
 - i. And telling the IRB about them?
- b. Integrate the OSTP response team with RDS to coordinate services
 - i. Understanding the limitations
 - ii. Gathering Notice of Awards from different organizations for analysis
- c. Must find a way to reach researchers before IRB and consent process - marketing strategies
 - i. Brand as re-consent form?
 - ii. Ethos/new common rule delayed changes till July, wait to work with them after changes.

28. Data storage

- a. Closely coordinate with other data storage groups on campus to understand the options and best make referrals/recommendations
 - i. Data Storage Council launching 1-26-2018

29. Data transfer

- a. Learn about data transfer resources available from MSI, know when to refer
- b. In relationship to embedded data management, data managers should partner with IT liaisons when consulting to help minimize need for data transfer.
 - i. For example, can data be stored in single server location and accessed from multiple sources? Data transfer issues may be able to addressed in workflow.

30. Data use agreements

- a. Need to learn more about the office of Unfunded Research Agreements in OVPR, as they consult on data use agreements.
- b. Clear guides for resources
 - i. DUAs for requesting existing data from an external source (IT, college - our AD for Research has signed ones for ICPSR, is UFRA right for this?)
 - ii. Writing DUAs for data the researcher/University produces to be shared to others (UFRA & libraries should be involved).
- c. Implement custom DUAs in DRUM to allow more flexible sharing (not from documented themes).
 - i. Flexible terms of use like dataverse. Credential login.
 - ii. Have a centralized manager for embargoed datasets (someone from SPH who reviews data requests)

31. Data Visualization

- a. Presentation or public facing data visualization services (journalism librarian role?)
 - i. Consultation/guidelines for best practices in what makes a "good" versus "bad" visualization, data literacy, data ethics
 - ii. Service that "transforms" graph into a more compelling visual design (something along the lines of what 538 does to the unsexy graphs used articles they cover)
 - 1. Reach out to IT, Univ relations, Communication specialists
- b. Provide access to data visualization tools, partner with others on campus that host visualization tools
 - i. Uspatial for GIS/story mapping
 - ii. AHC/EDMR for Tableau
 - iii. Graphic design (e.g, Photoshop, Illustrator)
- c. Communications

- d. Others?
- e. Develop expertise on the range of visualization tools for research data
- f. Visualization tool "Selector" could be a nice website page (LATIS gathered some data to make something like this a while ago, happy to share with someone who could update and implement)
 - i. Consult on how to use the tools, refer as needed to other offices as appropriate for specific/complex visualizations or tools

32. Database Development

- a. Develop guidelines to determine the "right size" database need for a given research project
 - i. When is a relational database better than a flat file?
 - ii. When to use a standalone database (such as Microsoft Access) or a hosted database?
 - iii. When is a "out of the box" tool better than designing a new database?
 - iv. Development of database access mechanisms (webpages, SQL queries, etc)
- b. Explore campus resources for hosting, developing, maintaining, and designing databases
 - i. Each step may be different, and not all geared towards non-tech researchers

33. Deidentification (Theme: Sensitive Data, see also DUA and DSA)

- a. Review DRUM collection policies for accepting sensitive data sets (e.g., whether or not to accept embargoed or "enclaved" data sets).
 - i. Explore potential levels of support for sensitive data in DRUM:
 - 1. More involved/customizable data use agreements that users must click through (and potentially provide name)
 - 2. Permanent embargos for data that requires approval (from author? From us?)
- b. Develop assistance or referrals for services to ensure the ethical and legal treatment of private or sensitive data (e.g., anonymization services). - what are services we should charge for vs do for free vs refer out?
 - i. Rather than libraries, share this info with <https://www.healthprivacy.umn.edu/>
- c. Provide researchers with training/templates specific to IRB and consent considerations when planning to share data.

34. Digital mapping: Create online tutorial or guide on resources for digital mapping for student classroom assignments.

35. Discovery and Access to Data [combined with "Data Registry" points]

- a. Create a "registry" or "data catalog" of data sets which are available to researchers (locally hosted/published and licensed/purchased) [within existing discovery/access framework/workflows, not as separate database, etc.]
 - i. Could just be a resource type within Alma/Primo then using the API / RSS push out to a separate interface (if desired)

- b. Maintain list of reliable data sources for reference and discovery.
 - i. Currently do this here: <https://www.lib.umn.edu/datamanagement/find>
 - ii. lib.umn.edu/data is another potential place.
 - iii. Library web page: How to Find Data
- c. Advertise more aggressively that the libraries can help researchers find data sets they can use.
- d. Website/landing page that unites the various data sources - map, government, text data, SCOPUS

36. DMPs Data Management Plan (DMP)

- a. Get SPA to alert us when a new grant is awarded and let us read the DMP so we can set up a meeting to make it actionable.
- b. Transform the DMP consultation service into a DMP review service that provides fast and timely feedback on draft plans.
 - i. Also set up system to ensure they get help following through with the DMP if funded - liaison follow ups, notifications from SPA about successful grant funding, etc.
- c. Update the Libraries' DMP template (originally written for NSF) to account for other agency requirements.
- d. Promote existing services for consulting with researchers on writing a DMP for federal grant applications. Expand this service to include a variety of non-federal funding agency requirements.
- e. Teach DMP workshops to general and targeted audiences (grant coordinators, departments, research groups, etc.) and varying promotion techniques to reach new audiences.
- f. Update and promote the UMN integration of the DMPTool.
 - i. Make clear that it needed to be restricted to 2-pages
 - ii. Needs more integration (e.g., storage options)
 - iii. Update what's currently there
- g. An important non-technology related aspect of this theme will be improving or maintaining our working relationships with OVPR, SPA, and grants coordinators and administrators. Working with these units and groups could help us reach researchers and give us the opportunity to better incorporate our tools and services into researcher workflows.
- h. Involve subject liaisons when consulting with faculty. Liaisons will be one of the best sources for consultation on the first point of contact for their department should be maintained.
 - i. Quick turnaround time at odds with this
 - ii. Concerned about response time / follow through

37. Documentation and Organization [add to data management theme]

- a. Assist researchers in creating data documentation to prepare data for eventual sharing and reuse.
 - i. Currently use the *readme.txt* template at <http://z.umn.edu/readme>
- b. Researchers need better tools to help them collect and document their data.

- i. More “flavors” of readme templates, etc.
 - ii. Custom metadata schemas development for better data collection.
- c. More awareness around automatic metadata collection by software
- d. Provide best practice on how to organize data to facilitate reuse. For example, adding descriptions, filenames, snapshots or “working” versions of the data.

38. DRUM

- a. Update policy and web submission form to highlight the unlimited upload limit(over 2GB per file).
- b. Issue open data badges to submitters. Highlight features of a record in a badge like way (see re3Data.org)
- c. Create a staging platform for manipulating DRUM files before upload while capturing consent from authors.
- d. Offer pre and post submission consultation.
- e. Track impact measures of data in DRUM (citations, etc.) Consider using a AltMetric package.\$
- f. Enhance usability/UX features of DRUM upload process
 - i. Automate file format transformation on ingest for certain files (MS Word to PDF).
 - ii. Allow drag and drop files
- g. Assess our data use policy adopted by the UMN Libraries. Is this working?
- h. Implement API or web service that allows users to analyze data (such as spatial tools like MapServer and ArcGIS Server) or serve archived data for alternative display, analysis, and visualization in another platform. (see data verse)

39. DRUM and big data

- a. Partner on building infrastructure internally that supports the data curation process (e.g., direct push to DRUM from MSI via Globus; Elevator to DRUM).
- b. Investigate a different platform that integrates compute into the storage for sharing and using big data

40. DRUM and Creative Commons Licenses

- a. Display clearly the license of a dataset even when no CC is selected.
- b. Make CC0 the default of all submissions (current is all rights reserved)
- c. Present users with instructions on selecting a CC and why it matters

41. DRUM and Discovery

- a. Include DRUM records in the catalog
- b. Include DRUM records in Experts
- c. Create a webpage with instructions on how to access DRUMs open API feed.

42. DRUM and DMP

- a. Review boilerplate language for using DRUM in DMPs and update

- b. Make an actionable timeline for implementing DRUM in DMP.
- c. Review UMN DMPs and follow-up with those who included DRUM.

43. DRUM and Publishing

- a. Allow users to version their own records in DRUM or to add publication metadata
 - i. Update permissions
 - ii. Include "Version/Update" Links next to Past Submissions.
- b. Engage with partners on campus to explore potential peer-review services for data.
 - i. Or reproducibility checks for manuscripts about to be submitted - does their code produce their results? Do the tables and reported numbers match?
- c. Build out workflow to easily enable authors to submit data upon submission of manuscript to a journal that requires data sharing
 - i. "Under review" status for data that indicates it could change to reflect reviewer feedback on manuscript/analysis (prevents need for versioning)
 - ii. Mediated access to data and code that enables reviewers/editors to access the data confidentially (without the author needing to field the request), but is not publicly available until manuscript acceptance.

44. DRUM: Data Repository and Curation Services

- a. Take advantage of linked open data technologies or networked research information (e.g., Open Folklore) in strategic areas of campus domain expertise.
- b. Research existing software and hardware and provide scalable solutions for visualization and virtual emulation services that might be incorporated into a data archiving platform.

45. ELNs

- a. Support electronic lab notebook and data documentation tools and software (e.g., LabArchives, Open Science Framework) through web and consultation services.
- b. License an ELN package such as LabArchives (UW, Cornell have experiences to share).
 - i. May not be a big need...groups either license a highly specialized tool for themselves or want something free, easy.
- c. Review report of Libraries ELN group and consult with authors about current need for support and possible roles/actions.
- d. Offer ELN training sessions (beyond the bootcamps).

46. Electronic Theses & Dissertations

- a. Work with Grad school/GSSP to create DMP requirements for all students
 - i. Consult with a librarian part of regular thesis planning
 - ii. Update manuscript guidance with instructions on how to deposit data to UDC as supp files for thesis record or as

seperate DRUM record.

- b. Formal collection development of significant data sets such as generated from student work (e.g., theses, dissertations) and university generated data to showcase student and department outputs.

47. GIS Data

- a. Create a landing page of all data services and include GIS discovery and creation services.
 - i. Incorporate into find data page (lib.umn.edu/data?)
- b. Recommend GIS data repositories and showcase BTAA Geoportal
- c. Include GIS data training on other data training advertisements.

48. Intellectual Property (IP)

- a. Provide workshops, learning objects and/or consultations on:
 - i. copyright, ethics, and data.
 - ii. data licensing, how to select.
 - iii. intellectual property rights.
 - iv. Policies

49. Long-Term Storage [*see: Data Preservation*]

50. Makerspaces

- a. Consult with Makerspace/Technology-Rich User Spaces Team (TRUST) co-leads about possible roles for Research Data Services Team (RDST).
 - i. E.g., publish 3D printer files created in makerspaces into DRUM or other public access repository
- b. Work with Makerspace leads to develop information and links to existing data services for websites.
- c. Include potential data services in process of asking students about future makerspace directions (to determine specific needs/roles), which will be determined by student users.
- d. Coordinate with other makerspace/technology spaces groups on campus (LATIS, CEHD Ed Tech group, Mechanical engineering/CSE, BBE, DASH) - role of TRUST group?

51. Marketing

- a. Develop a communications plan that:
 - i. Identifies the RDS services the libraries wants to promote
 - ii. Schedules timing and frequency of messages
 - iii. Plans specific modes (how) messages will be communicated
 - iv. Specifies who the audience will be for specific messages
 - v. Considers strategies for branding services

- b. RDS created posters & postcards that should be distributed

52. Metadata Standards Application/Development

- a. Develop a workflow for referring questions about metadata standards (to domain experts, controlled vocabulary experts, etc.). Use in-place ticketing systems (Service Now) to leverage in-house expertise and increase communication between people consulting on metadata for data sets, and those who can consult on controlled vocabulary adoption/creation, etc.
- b. Market services we can provide in this area, using more understandable terms [people don't know "ontologies" but they often want to develop a set of subject "tags"].

53. One Shot Workshops (Theme: Education and Training in Data Topics)

- a. Expand the graduate student-focused "Bootcamp" training in data management for a wide-range of disciplines. Engage with campus deans and related stakeholders on education issues surrounding data management. Ideas include:
 - i. CFANS - 200 students a year that need this training.
 - 1. Develop partnerships within the colleges to promote this training
 - ii. How to write a participant agreement (Note: LATIS has specific expertise and existing workshop series on these kinds of topics)
 - iii. Software-focused (e.g., "Using R for Reproducibility") - collaborate on existing [workshop series](#)
 - iv. Scaling is the problem - Opportunity costs/tradeoffs
 - v. Consider buying-out staff time, compensating presenters; data management education coordinator
- b. Target data management education to staff (e.g., project managers, lab managers) and post-docs.

54. Organizational structures

- a. Develop workflows for referring requests that would benefit from expertise distributed across units in the library. Leverage/coordinate in-place ticketing systems and increase communication across units about work being done and expertise available in-house.
 - i. How can other groups/units refer through tracking to libraries (e.g., IT also uses Service Now; DASH and DRUM use individual emails)
- b. Document the (sometimes overlapping) functions of various committees and groups in the Libraries and University to reduce reliance *solely* on individual connections.
 - i. Updating documentation; keeping connections if individuals leave
- c. Develop an internal-to-the-libraries skills inventory to assist in cross-divisional capacity building. (e.g., PIM citation managers; DASH did this maybe?)
 - i. Build into experts?
- d. Develop formal support mechanisms for cross-divisional referrals. (Admin support, see #1 too)

- e. Integrate individuals who are already doing work or research in an area into the groups and committees who take leadership on operationalizing the work at the Libraries level. Make sure that people who've been working on initiatives or built their research on areas have a voice in operationalizing. (Something analogous to libguides having a contact person for a specific content area?)
- f. Selection process for groups with leadership/decision-making roles happens at the AUL-level, and has a high chance of excluding people with expertise who might not already have high-level recognition of their work. Is there a way to open up a call for participation before the groups are finalized?
- g. Service ownership vs. silo formation vs. sustainability of the work if people leave.
- h. Organization-wide conversation about norms and practices around collaboration. Ensuring that collaboration is acknowledged and rewarded. (In particular in space of continuous appointment - devaluation of co-authored publications.)
- i. Acknowledging and recognizing cross-college collaborations on services in how they're promoted.
- j. Develop organizational structures (groups, units) to align related data efforts in the library (e.g., CIC GIS data platform, DRUM, etc.)
- k. Ensuring clear niches/objectives for each group created, as there seems to be many entwined efforts that may benefit from less division (RSC, RDS, OSTP, Reproducibility, etc)

55. Physical data management

- a. Offer consultations on physical data management, using principles of electronic data management
- b. Recommend museum artifact management software/best practices for individual/labs wanting to catalog their samples.
 - i. Connect individuals/labs with guidelines/best practices for selecting/implementing software, metadata schemas and connecting with others on campus doing similar work (case studies?)
 - ii. Connect to scanning labs (ASIOS) about management of physical samples after scanning

56. Planning

- a. Coordinate with other groups on campus to create visionary/strategic plan focused on data services for UMN as a whole
- b. Test/pilot new models of cross-collaboration, centralized work
- c. Policy development and Implementation
- d. Identify potential partners and collaborate with them (e.g., IRB, IT units) to coordinate services and policy implementation.

57. Project Close Out

- a. Offer specific guidance and consultation around project close-out, to help reduce data loss
 - i. RDS team created postcard on this topic
 - ii. Coordinate with storage council and IT units who are archiving data

- b. Work with graduate school to target graduate students as they are finishing their theses/dissertations
 - i. [Graduate Student Services and Progress](#) actually controls process and helps with ETD submissions, grad school sets policy but does not implement.

58. Referral [merge with Planning?]

- a. Develop a strong, pro-active referral network of existing data services on campus and create awareness of these cross-departmental services amongst services providers.
 - i. For example, OIT offers a variety of data storage solutions that researchers are often unaware of in library workshops.
 - ii. *UCCS subcommittee "Data Storage Council" charged with addressing storage education recently formed to help address storage issues.*
- b. Actively pull in other service providers during consultations
 - i. Example: Encourage communication between library and IT liaisons when researchers have archival/data storage questions

59. Reproducible Workflows / Research Reproducibility

- a. Identify and promote workflows that support research reproducibility tactics.
 - i. Eg. Land of 10,000 workflows
- b. Rebrand data management services to better reflect growing interest/awareness/motivation of reproducibility.
 - i. Market DRUM as tool for transparency/open science (not just an archive)
- c. Provide services to support reproducibility (e.g., Code review/debugging).
 - i. New web page: <https://www.lib.umn.edu/researchsupport/reproducibility>

60. Restricted sharing [merge with data preservation]

- a. Currently we refer out for restricted sharing options - we could either continue that practice [which should include more formalized training for DRUM/RDS staff on proper repository referrals] or ...
- b. Offer restricted sharing in (or parallel to) DRUM [cost/benefit analysis should be conducted to determine if feasible].
 - i. Learn from archives/special collections, what are their guidelines? (Rosetta; Globus/GEMS at MSI)

61. Scholarly identity

- a. Integrate idea of scholarly identity into data services.
 - i. E.g., faculty website that brings together scholarly work with other outputs like educational materials.
 - ii. Academia.edu has this new feature, google sites has some templates.
 - iii. How can we track citations/impacts from materials not in Scopus
 - iv. Links back to how open you are as a scholar

- v. Researchers want something with a .edu URL, something that implies trustworthiness
- b. Make data release a well-recognized intellectual and scholarly contribution (CVs, T&P)
 - i. Incorporate data sets into the Librarian CV template, and recommend ways that the P&R committee can take data into account.
 - ii. Encourage lib staff to deposit public data into DRUM
- c. Coordinate with RSC and ROI Committee

62. Staff Training

- a. Create a comprehensive training program for all interested staff with three levels of knowledge: introductory (best practices, awareness), intermediate (comfort with performing in person consultations), advanced (methods/data visualization techniques, curators, train-the-trainer).
- b. Research and document courses and training opportunities around campus and keep an up to date list for interested staff (stats courses, ethics workshops, etc.).
 - i. MM or FF emails. Identify who takes the lead on sending these updates.
- c. Maintain list of conferences and journals with RDS focus and highlight to all staff when relevant (staff web?)
- d. Develop and implement an educational program to ensure that staff are adequately trained to respond to the interest in research data management, building on past efforts of the Research Service Coordinators.
- e. Developing our library staff education program will give us the opportunity to take a lead in the development of data management education among research libraries. Offering training sessions as webinars (eg. ACRL and ARL) is one way for we might use share our expertise and build a community of education sharing among our peers.

63. State Agency Data

- a. Promote data management materials and services to state agencies and partner with them to help them build best practices. (Some library staff working on this; DNR approached about DMPs and other web materials).
- b. Potentially expand scope of DRUM to ingest state agency data.

64. Study design/research planning

- a. Coordinate with grants coordinators to advertise available grant opportunities and grant preparation/writing services across campus.
 - i. [Advertise grants workshops](#) as research realm
- b. Familiarize relevant library staff with IRB processes to facilitate consultation on IRB submissions, coordinating with other offices as applicable.
- c. Work with IRB to develop guidelines for submitting protocols that allow for sharing and archival of data
 - i. CITI training is now through UMN portal. Invitation to add modules to this training - data sharing, data security

- d. Know the partners the researchers will need to connect with across the research lifecycle and help them navigate it as they are planning their study (IT, college research offices, IRB, MSI, IRSA, etc)
- e. Coordinate and partner with research and statistical offices (IRSA, CTSI, UMII) to provide research study design consulting. Need to have better understanding about the language used here (scientific methods vs workflows).

65. Systematic Reviews

- a. Systematic reviews teams in HSL and R&L will handle most of inquiries and needs. DRUM will already provide services for data curation and sharing.
 - i. Add a “data type” for systematic review data/bibliographic
- b. SR teams will provide educational sessions for libraries staff and learning objects that will address some aspects of SRs. Most questions seem like reasonable extensions of the SR services.
 - i. Awareness of people to refer out to, and available tools (particularly free tools such as Rayyan)

66. Text mining

- a. *Note: It's unclear from many of these requests whether they represent outstanding needs or whether they needs were current met with the discussion and/or referral.*
- b. Advertise text mining as something that the DASH network can help provide
- c. Gather resources of various text mining activities and tools
 - i. Collecting of texts:
 - 1. Workflow to refer to someone to E-Resource Management department for help navigating licenses. Connect with acquisitions folks to purchase licenses.
 - a. Digital humanities librarian used to be in charge of purchasing text
 - 2. Advocate for metadata from publishers and bibliographic tags to indicate "scrapable" resources
 - 3. Gather information about available APIs to external databases/publishers and information about how to find/navigate these resources
 - ii. Analysis of texts:
 - 1. Publish tools for preliminary text analysis, such as Voyant (do this on DASH)
 - 2. Voyant server instance for the U
 - 3. Collaborate with campus and DASH partners to provide more heavy coding/statistical modeling for text analysis

67. User-Needs Assessment

- 1. Evaluate and better understand the data management needs of campus and propose solutions to be implemented by the Library.

2. Needs might be assessed through structured investigations (e.g., surveys, informal network, referrals) or by assessing current data management practices through observation, reviewing DMP proposals (see 2013 project) and literature reviews of actual data use and practice (Philip Herold's article).
68. Video/Image Data: Mention existing services around video and image data creation and hosting services in data services website.
69. Web hosting
 - a. Already provide some types of web hosting solutions for digital assets and collections; RDS could help promote these through website and raise awareness among libraries staff.
 - b. Develop specialized hosting/setting up/webUI services for databases. ([OIT does this](#)).
 - c. Develop best practices for archiving/publishing databases. (See DATABASE)
70. Web-scraping
 - a. Gather and publish resources around ethics of web-scraping
 - b. Consult on tools for web-scraping, collaborating with other partners on campus as needed.
 - c. Conversely, education about how you give away your personal data to third-party websites. (Me&23, Personal "data cleanse" from online, right to forget laws in other countries = scholarly identity management; terms of use, "online" visualization/analysis tools that requires upload) -> intersection with research ethics education

Appendix B: Charge for University Libraries Research Data Services Strategic Planning Task Force

Approved 11-7-2017

Background and purpose

The University of Minnesota Libraries has historically provided strong research data services to campus users, providing significant services in the early planning stages (e.g., DMP consultations, discovery/purchase of data sets for reuse) and later stages (e.g., repository and curation services for data), and our well-recognized web presence has reached thousands of users worldwide. Today, the growing volume of researcher-produced data on campus, combined with increased competition for funding, demands more aggressive data services, education, and intervention by the University Libraries and strategic partners on campus and beyond. Building from the research

data services retreat held for staff on July 10, 2017 the University Libraries will prioritize the research data services across the entire data lifecycle that will best meet user needs and can make the greatest use of library expertise and resources for the near-term.

Charge

The Research Data Services Strategic Planning Task Force is charged with developing a strategic plan for Libraries' research data services to be implemented over the next 1-3 years. The plan should identify and justify which research data services are near-term priority for the Libraries, roadmap how to move forward effectively providing those research data services (staffing, structure, resources, process flow etc.), and acknowledge research data service areas that are important but best managed by other units and campus services. The task force is expected to engage with non-library stakeholders at key points in their process. The scope and specific tasks are discussed in more detail below.

Scope

The Task force is intended to be focused primarily on:

- Data services oriented toward campus- and user-needs, and primarily focused on research applications of data as well as the use/integration of data and data management concepts into courses and curricula.
- Data management, discovery, and access is within scope only when defined by bullet one.
- Acquiring, curating, archiving, or providing access to data that has research value to our community, is at risk, but is not currently in scope for DRUM (purchased data, DNR and other state agency data, registry of data that resides elsewhere, use of existing data sets, licensing issues, etc.)
- Electronic lab notebooks (ELNs) and other research workflow documentation tools that describe data.

The scope of the Task Force does not currently include:

- Library data (gate counts, user-downloads, circulation data) and other administrative data that are managed on campus but are not widely accessible (e.g., peoplesoft data).
- Metadata, unless it directly pertains to a dataset.

Draft Work Plan and Anticipated Outcomes

The work of the task force will be broken into three phases to be carried out over four months.

Phase one: Information collection

Timeframe: Nov 1 - Dec 15, 2017

1. Define the need
 - a. Define existing data services in the context of the broader research lifecycle and for a variety of stakeholders (faculty, staff, grads, undergrads, public, etc.). Use themes from 2017 staff retreat as a starting point.
 - b. Gather, review, and summarize takeaways from past user needs assessments (see appendix) performed over the last three years. What additional evidence do we need? If any.
2. Identify new needs and characterize services currently offered but not yet formally recognized
 - a. Supplement current understanding of user-needs by interviewing and collecting user-experience stories from library staff (e.g., RDS team members) and/or interviews with researchers.
 - b. Visualize how needs fit existing Libraries data services.
 - i. Consider using use cases to demonstrate various scenarios.
3. Understand data service gaps
 - a. Review existing gap analysis of data services on campus (See appendix)
 - b. Review campus websites and (when needed) solicit feedback from campus service providers (email?, RDM iCoP?) on data service offerings across the data lifecycle. Ask how each unit is measuring researcher needs. What is their strategy/approach? Who is their audience? Who are their partners?
 - c. Update gap analysis of data services on campus

⇒ Phase one outcomes: Draft strategic plan that includes: definition of data services, summary of user needs (past and recent faculty personas), summary of how the library and other campus units meet those needs, and any identified gaps in data services.

Phase two: Stakeholder engagement around preliminary analysis

Timeframe: Dec 16-Jan 15, 2018

1. Prepare output from phase one for external engagement.
2. Present draft plan, that includes user-needs with defined/proposed levels of services, to non-library representatives (faculty, OIT, OVPR, data services providers across campus) to solicit feedback on these and other questions:
 - a. Do these needs reflect your experience with campus researchers?
 - b. Do these services (current) meet those needs?
 - c. Are there additional gaps in support?
 - d. Are there areas/opportunities for coordination in providing data services?

- e. Awareness and promotion is an issue on campus, how do we increase researcher knowledge of existing services?
- 3. Based on feedback, update the plan
- 4. If needed, recommend new services or ways to close the gaps.

⇒ Phase two output: Draft strategic plan vetted by non-library stakeholders.

Phase three: Finalize data services strategic plan

Timeframe: Jan 16-Feb 15, 2018

- 1. Vet plan with library administration (AULs).
- 2. To make the plan actionable, for each item identify:
 - a. Priority level: <specify>
 - b. Role-based action:
 - i. Library as Leader role: Identify roles and responsibilities for this work
 - 1. Charge existing groups with services
 - 2. Create new groups to carry out services
 - 3. Train existing staff/Hire new staff
 - ii. Library as Partner role: Identify people connections to this work, channels of communication, how to establish if ties do not exist.
 - iii. Library as Observer role: Monitor (who responsible) and recommend thresholds where action may be needed.
- 3. Format strategic plan for data services for broader sharing with staff and on campus.

⇒ Final Output: A vetted, fully-actionable strategic plan for Libraries Data Services.

Meeting schedule

This is a time-limited activity; the Task Force will begin November 2017. The work is anticipated to require 3-4 months, with a target completion date of February 2018. Because of the time limitation for this group, it is recommended that the Task Force meet, at minimum, on a bi-weekly schedule.

Communications

Group co-chairs will give monthly updates, via e-mail, to the sponsoring AULs and--as needed--to other identified stakeholders.

Web presence, reports and official documentation

Reports and official documentation will be kept within Google Drive. Documents will be shared with sponsors, and other identified stakeholders.

Sponsors

Claire Stewart, Associate University Librarian for Research & Learning, administrative lead

John Butler, Associate University Librarian for Data & Technology

Janet Bishop, Associate University Librarian for Content & Collections

Janice M. Jaguszewski, Associate University Librarian and Director, Health Sciences Libraries

Membership

Members:

- Lisa Johnston (chair), Research Data Management/Curation Lead (R&L)
- Philip Herold, Director for Agricultural, Biological, and Environmental Sciences and Interim Director for Physical Sciences and Engineering (R&L)
- Shannon Farrell, Natural Resources Librarian (R&L)
- Alicia Hofelich Mohr, Research Data Manager (CLA LATIS)
- Shanda Hunt, Public Health Library Liaison & Data Curation Specialist (HSL)
- Kelly Thompson, Metadata Analyst (D&T)

Stakeholders from outside the Libraries (an evolving list)

- Claudia Neuhauser, Associate Vice President for Research
- James Wilgenbusch, Minnesota Supercomputing Institute
- Rebecca Davies, Veterinary Diagnostic Laboratory
- Marc Hillmyer, Center for Sustainable Polymers
- Susan Telke, School of Public Health
- CFANS/CBS John Fieberg, Daniel Cariveau (bees), Greg Cuomo (ARD)
- Members of the Senate Research Committee - Ask Chair Bill Arnold when we have something more to share
- Lucy Fortson - Zooniverse
- Additional faculty and other key groups to be determined as process evolves
 - OVPR - Thierry Boudet tboudet@umn.edu
 - Council of Research Associate Deans

Related groups

- Research Data Management Services Team

Sponsor accountabilities

Sponsors will review progress of group and will give input on an “as needed basis” to questions from the co-chairs. Sponsors are also responsible for monitoring that the group is on track with its charge.

Chair accountabilities

The chair will be responsible for setting meeting times, documenting progress of task force and reporting back to the sponsors on a regular basis.

Group member accountabilities

There is an expectation that group members will be present at scheduled meetings and will participate in all activities directed towards completion of the goals listed within this charge.